Design and Optimization of Energy Systems

Prof. C. Balaji

Department of Mechanical Engineering

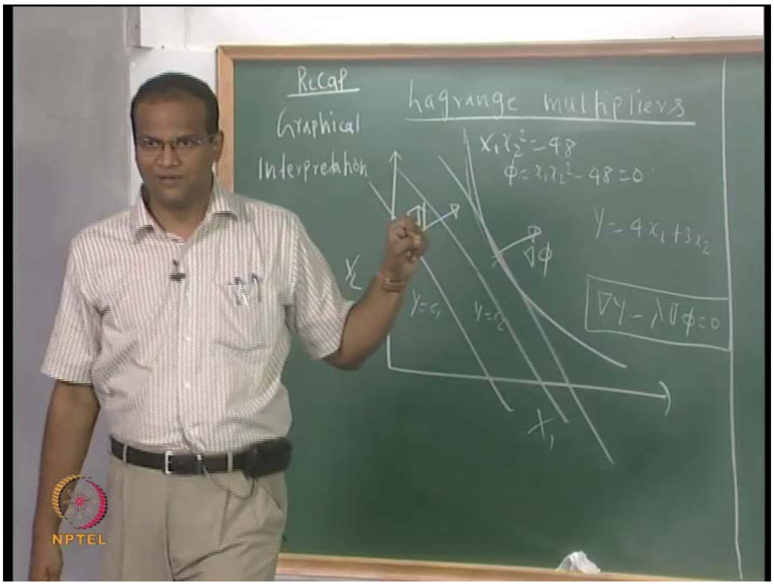
Indian Institute of Technology, Madras

Lecture No # 25

Mathematical proof of the Lagrange multiplier method

So, we were looking at Lagrange multipliers. Yesterday we looked at an interesting aspect of the Lagrange multipliers namely the graphical interpretation and the graphical visualization of the Lagrange multiplier method. In the first few minutes, I will repeat that point and then will go on to look at a mathematical proof of the Lagrange multiplier method and the economic significance of the multipliers and what are the second order necessary and sufficient conditions to establish whether the stationary point is a maximum or minimum or extremum. And we will wind up with some examples where we look at the second order necessary and sufficient conditions and try to figure out whether the optimum is indeed a maximum or minimum or inflection point.

(Refer Slide Time: 00:53)



So, if you look at this just a recap on the graphical illustration and a graphical interpretation. So, for a two variable problem this is x 1 and x 2 yesterday we drew a constraint like this, this was like I mean x 1 x 2 square is 48 or something, then so we drew Y is equal to c 1, Y is equal to c 2 and so on; Y is equal to 4 x 1 plus 3 x 2. So, these are all constant Y lines. So, this delta Y will be in this. So, this is phi equal to 0. So, this is a line this is a curve corresponding to phi equal to 0. Therefore, this will represent del phi. So, we are saying that del Y minus lambda del phi will be 0.

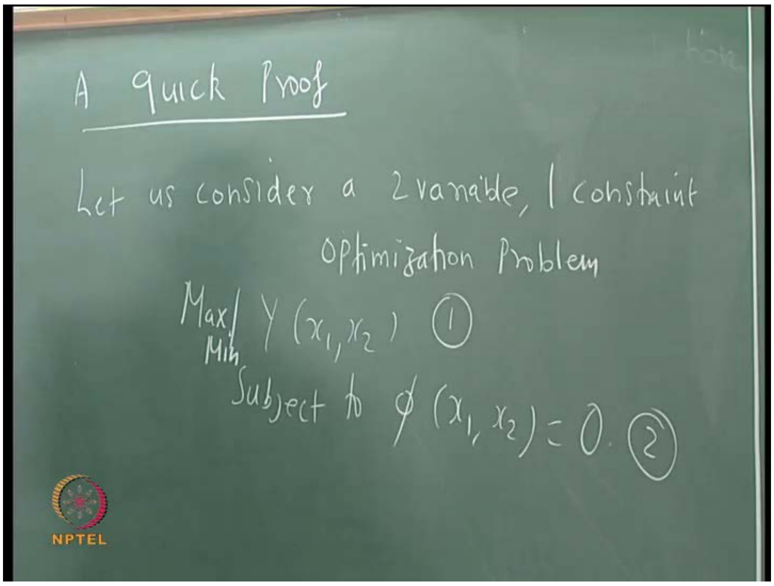
Yesterday there was a point whether suppose you move this Y is equal to constant into this, any problem Vinay? Suppose there was a question like whether if you move the Y is equal to constant line into this curve, what happens? I think Abhishek was trying to say, no, no it will satisfy the constraint; no it would not satisfy constraints, may be there will be some more points. See the constraints will be satisfied only if the point on this curve. So, if we go inside it does not mean that all the point, any point here will satisfy the constraints; no, it is x 1 x 2 square equal to 48.

Student: So, you have two other points.

Two other points may be but we cannot declare that whole region once you cross to the right of the tangent, the whole region becomes a feasible solution. Are you getting the point?

However, whatever we discussed was partly right. When we get into when we cross that point where you get the actual solution, there may be additional points which satisfy the constraint but which will definitely have a cost which is more than this, because Y is equal to constant is increasing in this direction. I think let us close the discussion on this.

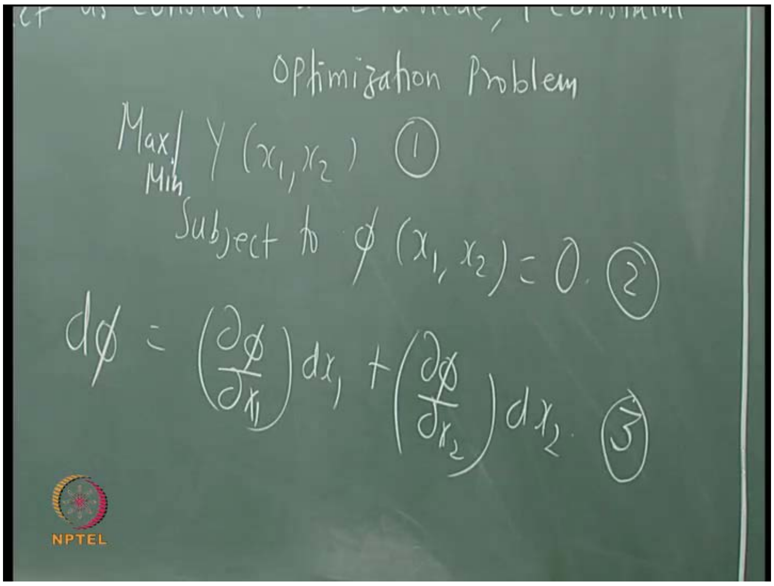
(Refer Slide Time: 03:21)



Now a quick proof, so the important point to remember is del Y and del phi the two vectors are collinear. Lambda, the introduction of lambda is a mathematical necessity because the magnitudes of del Y and del phi need not be the same. So you can have, some books will say del Y plus lambda del phi equal to 0, some people will say del Y minus, after all lambda is only a scalar, and then please remember that finally the point must be on the constraint. Therefore, lambda Y minus lambda del phi alone is not sufficient to solve the equations in conjunction with del phi is equal to 0.

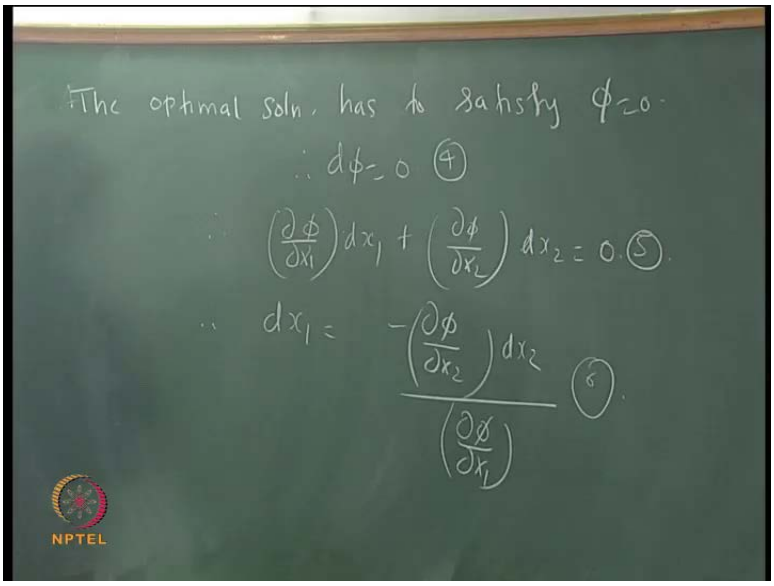
We will see a quick proof. Let us consider I mean this max or min Y maximize or minimize Y of x 1 x 2 subject to phi x 1 x 2 is equal to 0. This is a standard formulation of two variable one constraint optimization problem. So, we are seeking a solution to this.

(Refer Slide Time: 05:10)



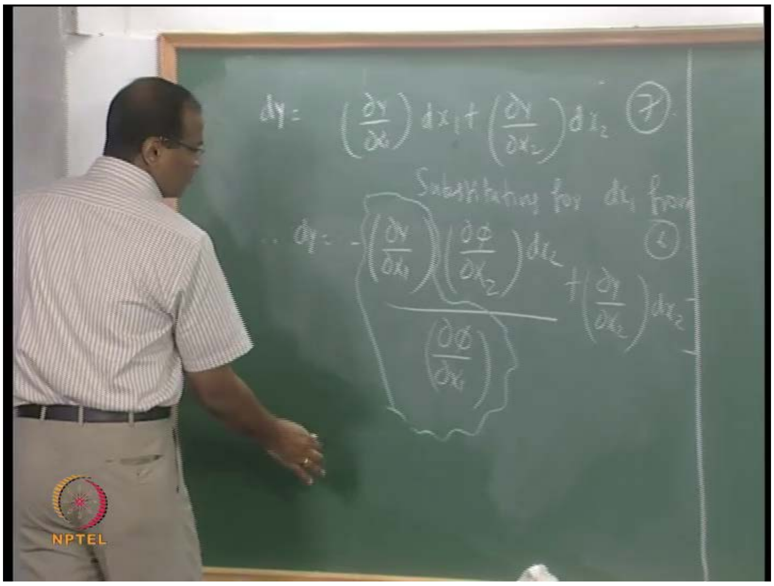
Now if you want to write out an expression for the differential d of phi, the d of phi can be written as, fine. When you are seeking a solution to the optimization problem the constraint has to be necessarily satisfied; is any d of phi allowed? d of phi has to be necessarily 0.

(Refer Slide Time: 05:57)



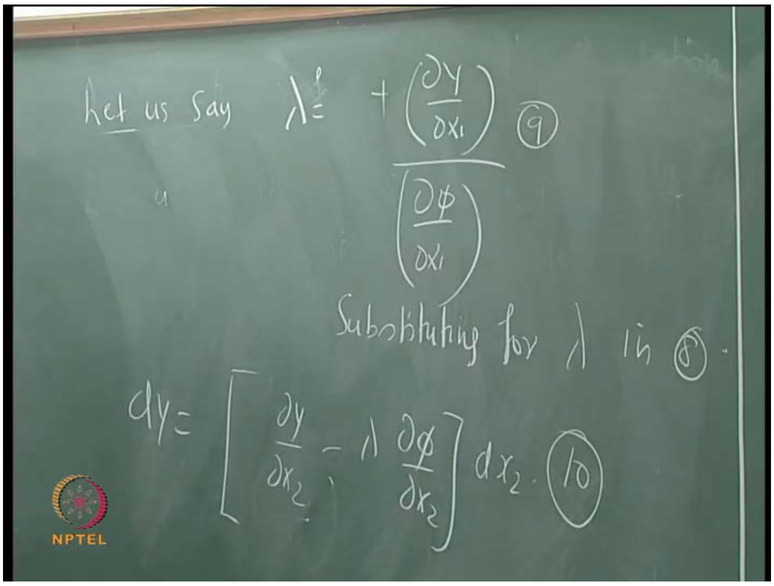
Okay any problem Vikram? Therefore, I am just doing some algebraic manipulation that is all. So, we got an expression for d x 1. Let us keep it. Now let us start working on d y, right. Shall I erase this?

(Refer Slide Time: 07:43)



Similarly, d y can be written as. Now it is possible for me to substitute for d x 1 in expression 7 from equation 6, correct. I come again; it is possible for me to substitute for d x 1 in equation 7 from the expression I have obtained in equation 6. Therefore minus dou phi by divided by plus. So, you can substitute for d x 1 from equation 6 and you can get this.

(Refer Slide Time: 09:56)



Now have a look at this and let us say lambda equal to minus. I am not violating anything, I can define. So, India is a free country so you can define whatever you want.

So, let us define lambda equal to. Now substituting for lambda, this should be 8, if you are seeking a solution to the optimization problem regardless of the value of delta x 1 and delta x 2; that means you have a point which could be a 1 a 2 or whatever; from a 1 a 2 you are changing that point by delta x 1 and delta x 2. So, regardless of the value of delta x 1 and delta x 2 if you are seeking the optimum, d y has to be necessarily equal to 0.

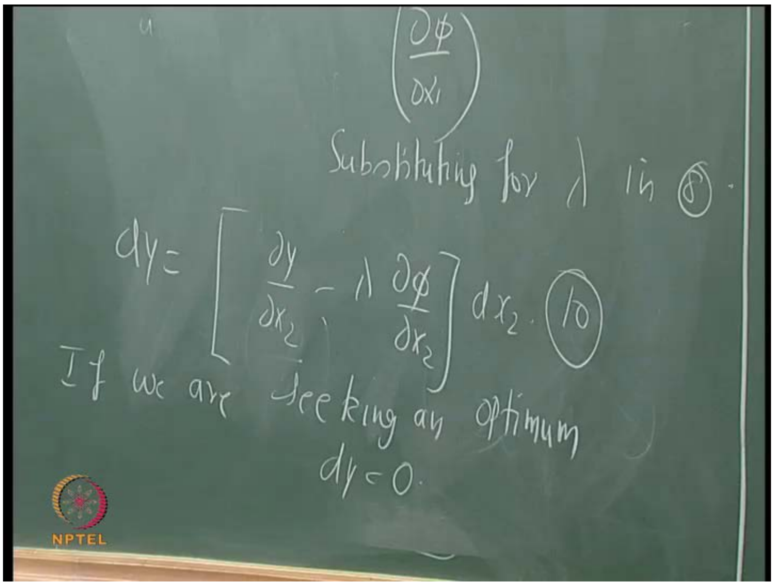
Student: Plus lambda.

Where? I have defined it.

Student: You have defined as minus lambda but its plus.

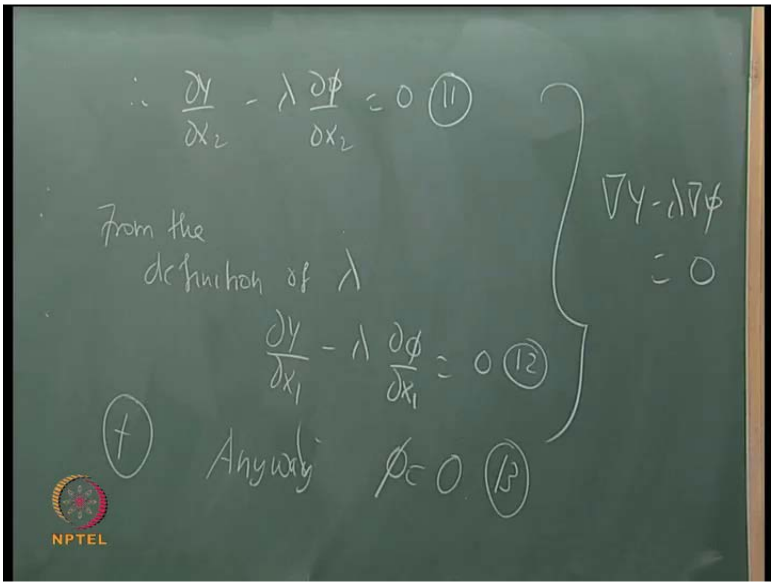
Then we will get the minus, okay. Does not matter; it is up to us but anyway we should not make mistakes on the board, right. What are they saying?

(Refer Slide Time: 12:02)



Yeah, so if you are seeking an optimum, therefore, the term within the bracket has to be 0. There is no point in making d x to 0; that is very silly argument, right.

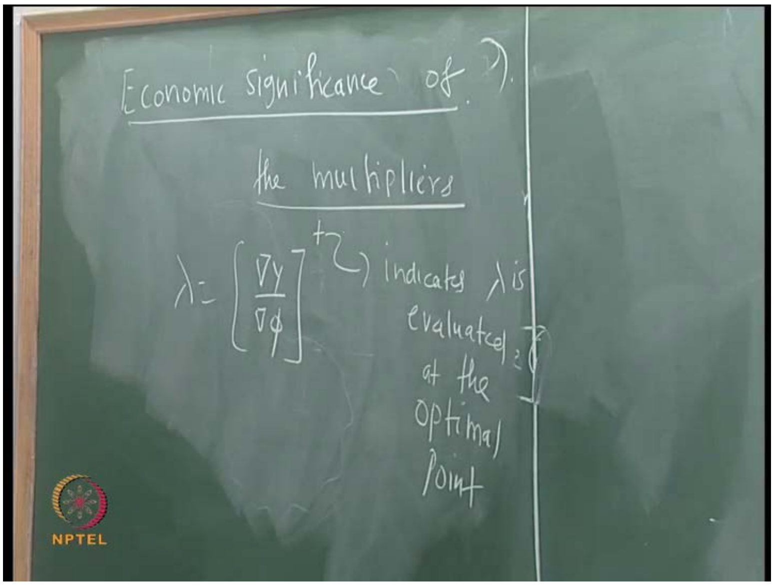
(Refer Slide Time: 12:29)



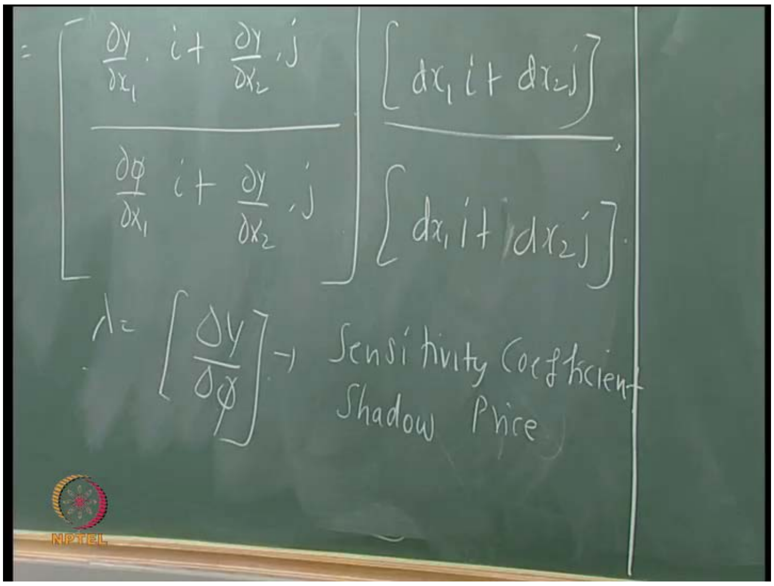
Therefore, from the definition of lambda d y by d x 1 minus lambda dou phi by dou x 1 is equal to 0 from the definition of lambda. Therefore, del Y minus lambda del phi equal to 0 for that two variable one constraint problem. You will get additional equations if we have more variables and more constraints. This is what we also try to interpret graphically, outside of all this plus plus plus anyway. So, equations 11 to 13 constitute a set of n plus m equations. If you have n variable m constraint optimization problem where m equality constraints are there, this can be solved and simultaneously get the values of x 1 to x n and lambda 1 to lambda n. This is a mathematical proof of the Lagrange's multiplier method, is that okay.

So, what we are trying to say is if you have a constraint phi equal to 0 and you are trying to move from the constraint d phi and you make the d phi equal to 0, and then get an expression for the differential of d y and make it 0 which means you are essentially trying to maximize Y or minimize Y subject to the condition that p is equal to 0. So, mathematically this making d phi is equal to 0 and d y equal to 0 is analogous to solving the set of equations which turnout to be like this. Therefore, this is essentially if you are solving this, you are essentially solving that, alright fine.

(Refer Slide Time: 15:10)



(Refer Slide Time: 16:02)



So, let us look at the economic significance of the multipliers. So, lambda this plus indicates at the optimal point, right. Vectorial form I can write d y by d x 1 i plus I can actually multiply by, what does it give you? d y d x 1 i d x i plus d y d x 2 basically it gives you, okay. I just change the inverted triangle to the normal triangle. Therefore, lambda gives the ratio of the change in the objective function to the change in the constraint. This is also called, so this is actually the sensitivity coefficient. In operations research it is also called the shadow price.

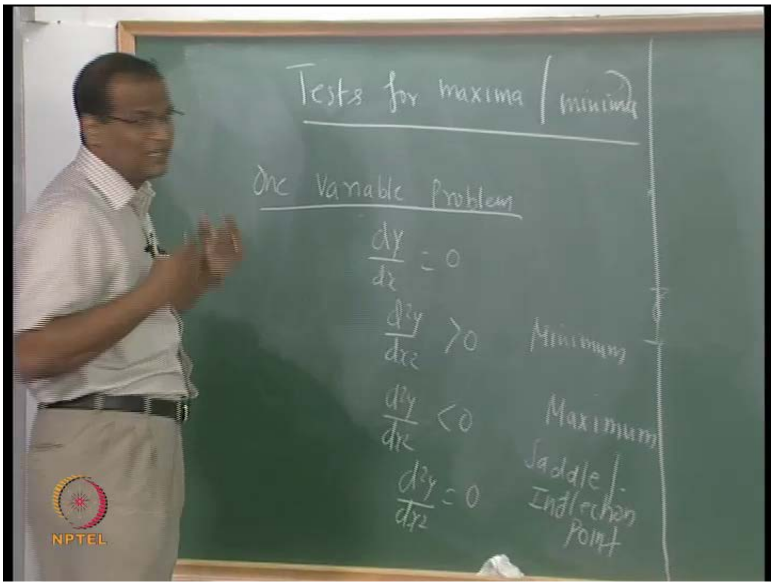
What is the meaning of the shadow price? Chalk piece. Let us consider a problem. So, there is a company which is making furniture. He is very much worried about the chalk piece, let it be there know. The soldier should have some blood, right? So, what is the shadow price? Let us look at a furniture company which makes only two types of products table and chair. So, certain amount of wood is required for making one chair, certain amount of wood is required for making one table, certain amount of labor is required for making the chair, certain amount of labor is required for making the table. The cost price of the chair will be smaller compared to the cost price of the table. Therefore, the profit will be some c 1 x 1 into c 2 x 2 where c 1 is basically the?

Student: Profit per chair.

Profit per chair and c 2 is a profit per table. You want to maximize the profit subject to the condition that there is a finite amount of labor and there is a finite amount of material available.

So, this Lagrange multiplier basically talks about. So, in this problem for the furniture problem under consideration, so the two constraints are you have the constraints on the time available the labor and the material and there is a profit which is Y. So, if more wood is made available to the furniture company or if more labor is available to the furniture company, how will the objective functions change? So, it is a shadow price. So, if you are pumping more resources, what will be the profit? It is a shadow because it is not done yet. If you put in those resources according to all these formulations you have, it should give results in so much change in the profit. That is why it is called the shadow price, alright. Now let us go to some trickier mathematical considerations; we have to look at the higher order, second order necessary and sufficient conditions.

(Refer Slide Time: 20:18)

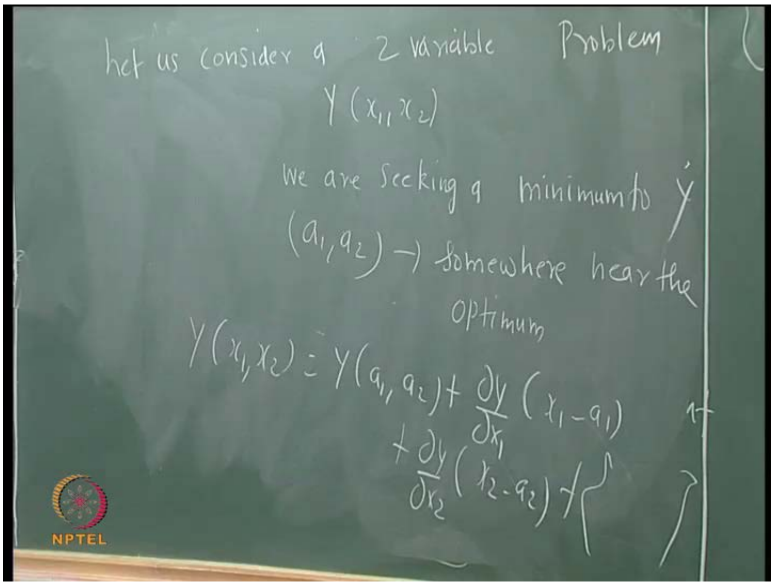


So, tests for maxima/minima one variable problem. d y by d x equal to 0, agree? It is called a saddle or inflection point; that means you do not know the second order test is insufficient or the function is moving very gently over there because it is very good for us. Mathematicians would like to get a global optimum or global, they want precisely a value of the x 1 to x n at which Y will become maximum or minimum. But this is dreaded by engineers because all the variables some measurement is associated with them; they are all subjected to errors.

So, I would love to have an optimum of Y which is not very sensitive to the values of excess at the optimum; that is the engineer's requirement. I do not want something going like this, because I cannot get that Reynolds number 3444, temperature 64.6 degrees and it will not come. So, many variations will be; ambient temperature will change, morning to evening it will change, seasons to monthly it will change, seasonal variation. So, mathematics is only a tool, we should not get carried away; everything cannot be solved and actually when something is working, so many things are there are so many imponderables. You do not have control over many things. Therefore, we should have some objective function which gives you some breathing space where if a value of Y is 100; it should be possible for you to get between 95 and 100 for a reasonable range of the variables, then it is fine.

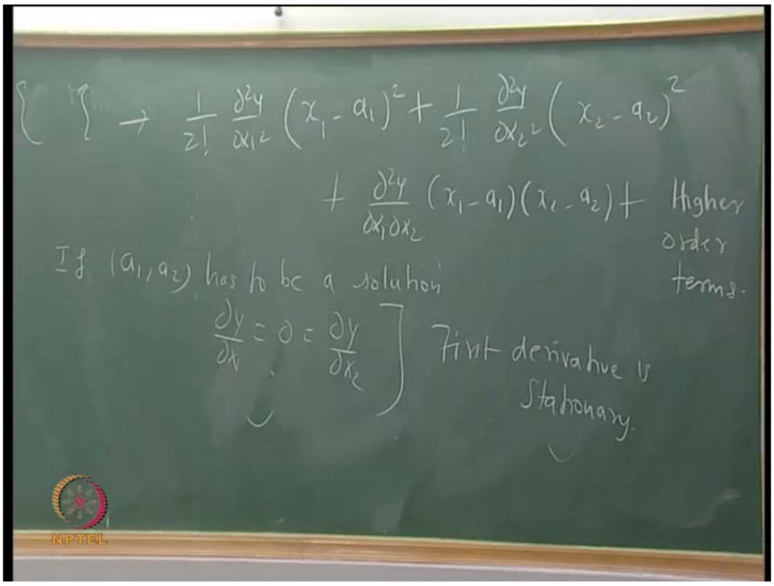
So, I will say that any of those solutions which gives Y 95 and above, it is alright. So, the sensitivity is very very important, alright. Now what about the test for maximum or minimum when more than one variable is encountered which is invariably the case in optimization, because one variable problem is the simple high school stuff; we cannot reduce all the problems to one variable problem. Also yesterday, we burnt our fingers with an example simple parabola problem where reducing it to one variable problem; we did not get the solution. Therefore, often times we encounter a multivariable problem; it should be possible for us to come out with test to determine the maxima or minima.

(Refer Slide Time: 23:38)



Let us consider a two variable problem. Y is the function of x 1, x 2. We want to minimize Y. We are seeking a minimum 2 Y or minimum of Y, whatever. Let us say that a 1, a 2 is a point somewhere near the optimum or is an optimum itself, but we want to test, right. Can you expand Y around a 1, a 2 using Taylor series? It is possible for us; you can expand Y of x 1, x 2 around a 1, a 2 using a Taylor series. This will be equal to plus, what is that?

(Refer Slide Time: 25:14)

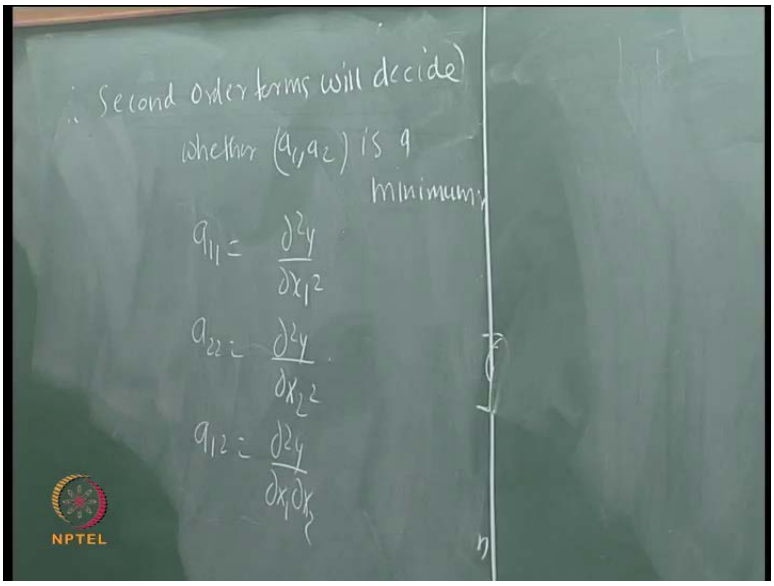


1 by 2 factorial into x 1 minus whole square plus plus higher order terms, let us the higher order term or assuming the higher order terms do not contribute much, okay. We should start examining, how do we establish the test of how do we come up with the procedure; let us determine whether that a 1, a 2 is really minimum or not. So, what is the first step? If a 1, a 2 has to be a solution d y by d x 1 must be equal to d y by d x 2, not a 1, a 2, correct. So, first order because otherwise you can simply change; if d y by d x is a large value you can simply move the point from a 1, a 2 some nearby point and then increase a function. If you increase a function beyond this a 1, a 2 plus; that means already or you can decrease the function depending on whether you are seeking a maximum or minimum, are you getting the point. Therefore, a 1, a 2 will no longer be a solution to the problem.

Therefore, first order conditions becoming the first derivative becoming 0 is a mandatory requirement. This is the first derivative is stationary. It is a r Y, right, e r Y is something else, e e r Y is paper and pencil. Unfortunately, Microsoft word will not correct because both are acceptable spelling. It has not become intelligent enough to figure out the context, right, I erased it anyway. So, what is the story now? Higher order terms are neglected first derivative become 0. Therefore, the second order terms have to be positive or negative or 0 what is it? If a 1, a 2 has to be solution, if you move from a 1 a 2 any perturbation in a 1 a 2 will result in a value of Y which is more than Y at a 1, a 2.

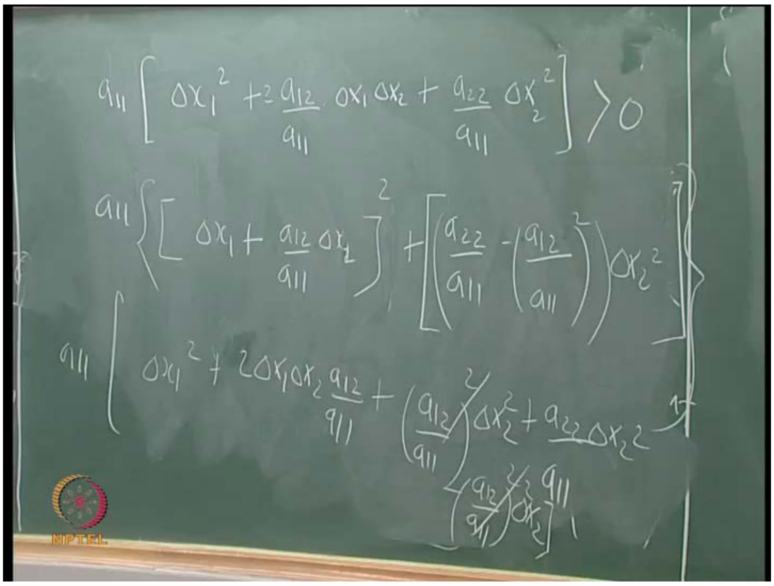
Therefore it is enough for us to prove that the second order terms result in a positive quantity. I am seeking a minimum only for a minimum, correct. For a minimum any perturbation in a 1, a 2, I mean any perturbation in x 1, x 2 around a 1, a 2 will result in a value of Y which is more than the value of Y at a 1, a 2. Therefore, Y at a 1, a 2 is indeed the minimum. In order to prove that the second order terms must be you have to prove that the second order terms or find out conditions when those second order terms together will become positive, is that okay.

(Refer Slide Time: 29:47)



Now in order to make the working easier, I will introduce some notation a 1 a 1 a 11 is dou square Y, a 22, a 12. The order of the differentiation does not matter, dou square Y dou x 1 dou x 2 is same as dou square Y dou x 2 dou x 1. So, can I look at the second order terms now?

(Refer Slide Time: 30:52)



So, shall we start with a 11 plus plus must be greater than 0, right. If this condition is satisfied you are getting a minimum, regardless of the value of delta x 1 delta x 2, right. The story is clear now. Now how do we get conditions for this? A very crude way is keep on changing delta x 1 delta x 2 and find this out; I mean that is very unimaginative, right. You keep changing delta x 1 and x 2 and find out what happens to this, you take twenty values of delta x 1, twenty values of delta x 2, write a small program. You did it for delta x 1 delta x 2; suppose you have got for 20 variables, 100 variables, 200 variables, there should be something which is mathematically more precise or which will help you get this. Now can you tell me? Can you do some mathematical manipulation of this? Up to this it is clear?

Student: Find out x 1 by solving a quadratic equation.

How do you do that?

Student: Take x square in the form of x square plus a 1 a x plus c equal to 0.

No, regardless to the value of delta x 1 and delta x 2, in fact you know that it is coefficients a 11, a 12 it will be these coefficients which ultimately decide the fate of this expression, right.

Student: Square by delta x square.

Fine, you start doing; I am going to work out, you start. Now I established the condition; I have told you that this is the starting point, there are many ways of doing it. Please start doing, yeah, you can rearrange that 4 and other things, the 4 is a problem; that is why I took this 2 out and I did something. I already multiplied, right. No, no I should get a 2 here, right. Yeah, yeah it is okay. It is a 2 there, sorry I mean it is so obvious that I missed the 2. Okay what you do now?

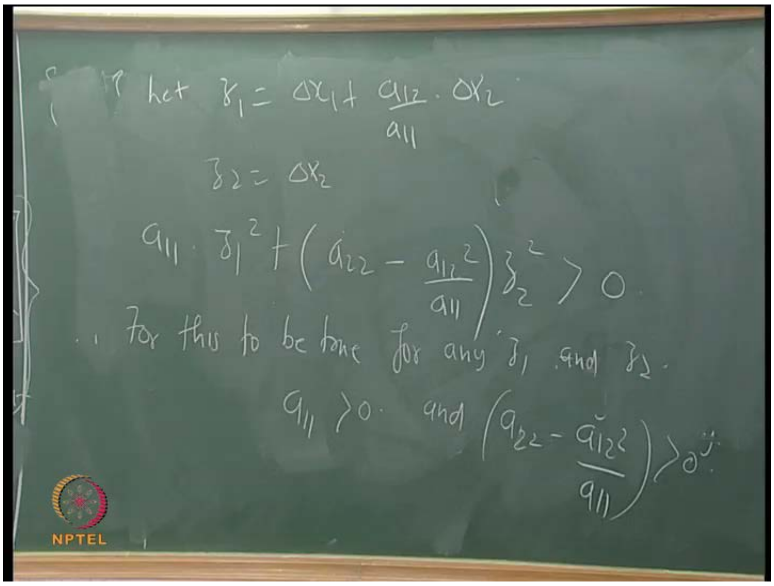
Student: delta x 1 delta x 2 into a 1.

Yeah, what did you do?

Student: Discriminant is less than 1.

Discriminant, yeah that condition can be satisfied. But I will work it out so as to save time. What I did was, can you write like this? a 11 into delta x 1 plus a 12 delta x 2 whole square plus, is that okay. No, a 11 will be common for everything, just tell me if it is okay; the a 11 will be common for everything, correct. What I have done now is this will be a 11 into delta x 1 square plus 2 delta x 1 delta x 2 a 12 by a 11. So, the first two terms you get. You will get an additional square term which is a 12 by a 11 square delta x 2 but I have put the square minus term here which will cancel, right. So, this will be plus a 22 by a 11 delta x 2 square which is coming here and I am subtracting this again a 12 by a 11 square delta x 2 square, right. So, adding and subtracting. This is a very roundabout way of doing but that is all right.

(Refer Slide Time: 36:08)



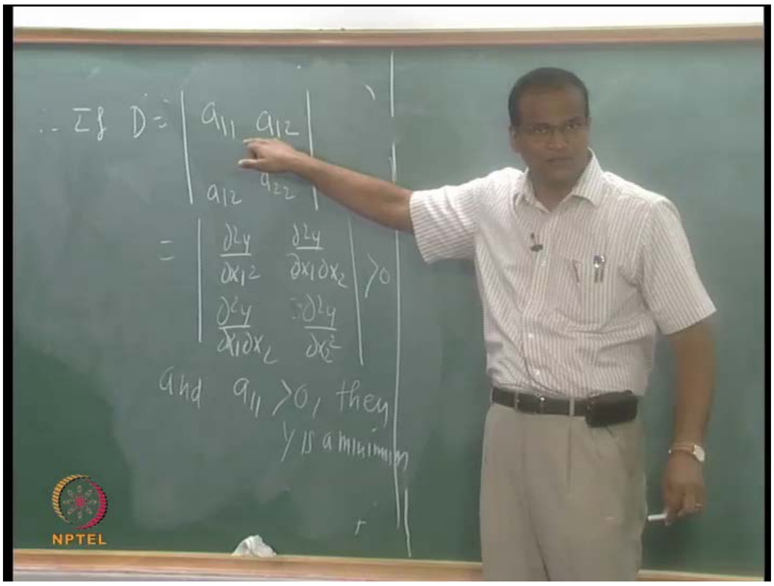
So, let z 1 be delta x 1 plus a 1 and z 2 equal to delta x 2. Therefore, the resulting expression is a 11 z 1 square plus can you tell me? What happened? a 22 minus, yeah can you complete it?

Student: a 12 square by a 11 square.

a 12 square by a 11 must be greater than 0. I did not do any magic; I just did algebraic manipulation that is it. Of course, when I am very liberally multiplying and dividing by a 11, the intrinsic assumption is a 11 is not equal to 0. If a 11 is 0 we have to go home, we have to pack. That is inflection point, already a 11 if it is 0 means where there is no scope for us to play, that should not be 0. Even for a single variable if a 11 that is d square by d x square equal to 0 we are in trouble.

Now if for all values of z 1 and z 2 this has to be true, individually the terms have to be positive. Do not try to find out combinations wherein for some orbit value of z 1 and orbit value of z 2, this combination is such that one is positive that would not work. For any value of z 1 and z 2, if the left hand side has to be greater than 0, a 11 and the term within the brackets have to be both greater than 0. Therefore, for this to be true for any z 1 and z 2, a 11 must be greater than 0 and a 22, okay.

(Refer Slide Time: 39:03)

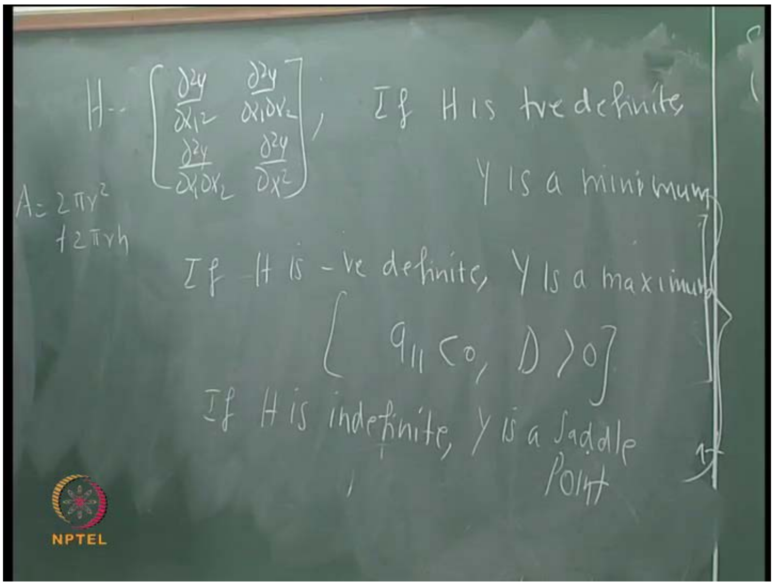


Therefore, if D is a matrix of is also equal to. If D is greater than 0 and a 11 is greater than 0, then Y is a?

Student: Minimum.

Y is a minimum. So, the D happens to be the determinant of the matrix of partial derivatives. I come again; D happens to be the determinant of the matrix containing the partial derivative, partial derivatives of second order.

(Refer Slide Time: 40:19)



Therefore, if you have the Hessian matrix which is defined by, if the determinant is greater than 0 and a 11 is greater than 0, H is called a positive definite matrix, okay. So, if H is positive definite, Y is a minimum. If H is negative definite, Y is a maximum; that is a 11 less than 0 d greater than 0. If H is indefinite, Y is a saddle point; that means it is an inflection point we do not know. Immediately expect a question, what happens if D is equal to 0? If D equal to 0, then it is a degenerate critical point, the Hessian test is inconclusive. So, very rarely you can set mathematical equations which will lead to D equal to zero, but in most engineering problems this will not happen.

In fact, for most engineering problems without doing the test of the Hessian matrix you will be in a position to decide whether the resulting optimum is a maximum or minimum. However, to be threateningly formal and that to be in order to be clinically correct, you can go for the testing with the hessian matrix and be sure that the final extremum you obtain is really a minimum or a maximum. There is one more way of looking at it; that is you have the Hessian matrix of the partial derivatives of the second order, if you find the Eigen values of the hessian matrix; if all the Eigen values are positive, then it is a positive definite matrix. If all the Eigen values are negative, it is a negative definite matrix. If some values are greater than 0 and some values are less than 0, it is indefinite. So, it is possible.

So, once you evaluate, for example, the area of the cylindrical storage tank 2 pi r square plus 2 pi r h. So, dou A by dou r will be 4 pi r plus 2 pi, dou square a by dou r square will be 4 pi. So, the first term will be 4 pi; you can substitute each of this term. Then, you can multiply and take it minus lambda of i; where i is an identity matrix. Evaluate the lambda and then find out whether the Eigen values are positive or negative or some are positive or some are negative or you just see whether a 11 is greater than 0, a 1 and a 11 is less than 0 and take this determinant and figure out whether it is a maximum or minimum.

Now it may look hazy or nebulous. So, in the next class whatever problems we have considered, 1 of out of those 3 of 4 problems we have solved thus far, we will take 1 or 2 and establish that whatever we obtained was indeed a minimum or maximum. For example, the cylindrical storage problem you could convert it to one variable problem and then we mathematically establish that it is a minimum. Now treating it as a two variable problem if you are able to establish that it is still okay, then the results are self-consistent or they are all right, fine.

So, if you want more then you will have to read upon some advanced mathematics books Kreyszig. For example in Kreyszig lot of discussion is there about Hessian matrix and all that. So, the Hessian matrix is a matrix of partial derivatives. We take the determinant and find and determinant has to be greater than 0, okay. If the determinant is less than 0, it is a saddle point; determinant has to be greater than 0. If determinant is equal to 0, which will never happen; if determinant is equal to 0, then this test is inconclusive. So, once you get the determinant greater than 0, then look for a 11. If a 11 is less than 0, it is a max; if a 11 is greater than 0, it is a minimum. So, we will stop.